

Parametrized Positivity Preserving Flux Limiters for the High Order Finite Difference WENO Scheme Solving Compressible Euler Equations

Tao Xiong ¹ Jing-Mei Qiu ² Zhengfu Xu ³

Abstract

In this paper, we develop parametrized positivity satisfying flux limiters for the high order finite difference Runge-Kutta weighted essentially non-oscillatory (WENO) scheme solving compressible Euler equations to maintain positive density and pressure. Negative density and pressure, which often leads to simulation blow-ups or nonphysical solutions, emerges from many high resolution computations in some extreme cases. The methodology we propose in this paper is a nontrivial generalization of the parametrized maximum principle preserving flux limiters for high order finite difference schemes solving scalar hyperbolic conservation laws [22, 10, 20]. To preserve the maximum principle, the high order flux is limited towards a first order monotone flux, where the limiting procedures are designed by decoupling linear maximum principle constraints. High order schemes with such flux limiters are shown to preserve the high order accuracy via local truncation error analysis and by extensive numerical experiments with mild CFL constraints. The parametrized flux limiting approach is generalized to the Euler system to preserve the positivity of density and pressure of numerical solutions via decoupling some nonlinear constraints. Compared with existing high order positivity preserving approaches [24, 26, 25], our proposed algorithm is positivity preserving by the design; it is computationally efficient and maintains high order spatial and temporal accuracy in our extensive numerical tests. Numerical tests are performed to demonstrate the efficiency and effectiveness of the proposed new algorithm.

Keywords: Compressible Euler equations; Positivity preserving; Parametrized flux limiters; High order finite difference method; WENO reconstruction

¹Department of Mathematics, University of Houston, Houston, 77204. E-mail: txiong@math.uh.edu.

²Department of Mathematics, University of Houston, Houston, 77204. E-mail: jingqiu@math.uh.edu. The first and second authors are supported by Air Force Office of Scientific Computing YIP grant FA9550-12-0318, NSF grant DMS-1217008.

³Department of Mathematical Science, Michigan Technological University, Houghton, 49931. E-mail: zhengfux@mtu.edu. Supported by NSF grant DMS-1316662.

1 Introduction

The success of the high order essentially non-oscillatory (ENO) or weight ENO (WENO) methods solving hyperbolic conservation laws has been well documented in the literature [6, 15, 12, 9] and the references therein. At the heart of the high order ENO/WENO schemes solving hyperbolic problem is the robustness, namely stability in the sense of suppressing spurious oscillations around discontinuities. The application of the high order finite difference, finite volume ENO/WENO methods to hyperbolic systems [15, 9], such as the compressible Euler equations

$$\begin{pmatrix} \rho \\ \rho u \\ E \end{pmatrix}_t + \begin{pmatrix} \rho u \\ \rho u + P \\ (E + P)u \end{pmatrix}_x = 0, \quad (1.1)$$

achieves the goal of suppressing oscillations when discontinuous solution emerges during the time evolution. However, in the extreme case, such as high Mach flow simulation, a slightly different (although equally important) problem is that the high order schemes that we are using might produce solutions with negative density and pressure, which leads to an ill-posed problem, often seen as blow-up of the numerical simulation. The failure of preserving positive density and pressure by the above mentioned schemes in such circumstance pose tremendous difficulty of applying high order schemes to some of the challenging simulations in practice.

In the earlier work, see [3, 11, 13] and references included, much attention has been paid to the positivity preservation of schemes up to second order. It wasn't until the recent work by Zhang & Shu [23] that arbitrarily high order finite volume WENO and discontinuous Galerkin methods are designed to preserve positivity. The method proposed in [23] is a successful generalization of their earlier work on the maximum principle preserving (MPP) computations of scalar conservation laws, see [24]. Their approach relies on limiting the reconstructed polynomials (finite volume WENO) or representing polynomials (discontinuous Galerkin) around cell averages to be MPP. The positivity preserving (PP) finite volume WENO scheme and DG scheme by Zhang & Shu can be proved to have the designed arbitrary high order accuracy when equipped with proper CFL number. In the later work by the authors [26], a PP finite difference WENO method is presented when the density and pressure is strictly greater than a fixed positive constant. In [8], a flux cut-off limiter method is applied to the high order finite difference WENO method to ensure positive density and pressure.

In this paper, we continue along the line of research on the parametrized flux limiters proposed in [22, 10, 20] for high order ENO/WENO methods solving a scalar hyperbolic

conservation law

$$u_t + f(u)_x = 0 \quad (1.2)$$

subject to the initial condition $u(x, 0) = u_0(x)$. For this particular family of equations, the solution satisfies a strict maximum principle

$$u_m \leq u(x, t) \leq u_M \quad \text{if} \quad u_m \leq u_0(x) \leq u_M. \quad (1.3)$$

The idea of the parametrized flux limiters for general conservative scheme solving scalar conservation laws is to modify high order numerical fluxes to enforce the discrete maximum principle for the updated solution. In general, a conservative high order scheme with explicit multi-stage Runge-Kutta (RK) time integration for (1.2) can be written as

$$u_j^{n+1} = u_j^n - \frac{\Delta t}{\Delta x} (\hat{H}_{j+\frac{1}{2}}^{rk} - \hat{H}_{j-\frac{1}{2}}^{rk}), \quad (1.4)$$

where $\hat{H}_{j\pm\frac{1}{2}}^{rk}$ are the corresponding fluxes at the final stage of RK methods. The MPP properties of high order schemes are realized by taking a convex combination of a high order flux $\hat{H}_{j+\frac{1}{2}}^{rk}$ and a first order monotone flux $\hat{h}_{j+\frac{1}{2}}$: $\tilde{H}_{j+\frac{1}{2}}^{rk} = \hat{h}_{j+\frac{1}{2}} + \theta_{j+\frac{1}{2}}(\hat{H}_{j+\frac{1}{2}}^{rk} - \hat{h}_{j+\frac{1}{2}})$, with $\theta_{j+\frac{1}{2}} \in [0, 1]$. The limiting parameters $\theta_{j+\frac{1}{2}}$, which measure the change of numerical fluxes, can be found out through decoupling the following MPP constraints that are linear with respect to $\theta_{j\pm\frac{1}{2}}$,

$$u_m \leq u_j^{n+1} = u_j^n - \frac{\Delta t}{\Delta x} (\tilde{H}_{j+\frac{1}{2}}^{rk} - \tilde{H}_{j-\frac{1}{2}}^{rk}) \leq u_M. \quad (1.5)$$

The similar idea is utilized in this paper in the sense of making sufficient modification of the high order numerical fluxes to ensure that the updated density and pressure are positive. When such parametrized flux limiters are generalized to preserve the positivity of density and pressure of numerical solutions for Euler equations with source terms, there are several new challenges. One of the main difficulties is that the linear MPP constraint (1.5) becomes nonlinear for positivity preservation of pressure, which has nonlinear dependence on the density, momentum and energy. We address such challenges by decoupling the nonlinear PP constraint for a ‘convex set’ of the limiting parameters. The proposed approach provides a sufficient condition for preserving positive pressure. The presence of the source term can also be conveniently handled in the parametrized flux limiting framework. Notice that we only require positivity preservation for the solutions at the final stage of RK method for the sake of preserving the designed high order temporal accuracy. If there are negative density and pressure in intermediate stages of the RK method, the speed of sound is computed by $c = \sqrt{\gamma \frac{|p|}{|\rho|}}$.

Our approach is similar to those very early discussions of the flux limiting approach [1, 2, 4, 17, 16] for the purpose of achieving a total variation diminishing (TVD) property, which is a much stronger stability requirement than the maximum principle. The schemes are expected to be TVD, therefore, most of the schemes are at most of second order accurate. To distinguish our work from others' in the context of designing arbitrarily high order schemes, we would like to point out that the method we are proposing only involves the modification of high order numerical fluxes. Another critical difference is that the parametrized flux limiters are only applied to the final stage of the multi-stage RK methods. These new features are designed to produce numerical solutions with positive density and pressure, while allowing for relatively large CFL numbers without sacrificing accuracy in our extensive numerical tests. The proposed method is essentially different from those by Zhang & Shu [26], in which the PP property is realized only with fine enough numerical meshes, when the density and the pressure is extremely close to 0. The flux limiting method we are proposing is also different from the flux cut-off method by Hu [8], whose approach demands significantly reduced CFL for accuracy as illustrated in their analysis and numerical tests. However, the proof of maintaining high order accuracy when the PP flux limiters are applied to the finite difference WENO method solving the Euler system is very difficult. In this paper, we rely on numerical observations to demonstrate the maintenance of high order accuracy. A rigorous proof of that the MPP flux limiters modify the original high order flux with up to third order accuracy for general nonlinear scalar cases is provided in [20] and that with up to fourth order accuracy for linear advection equations is provided in [21].

The paper is organized as follows. In Section 2, we give a brief review of the parametrized MPP flux limiters for high order conservative schemes solving (1.2). We then generalize the MPP flux limiters to a scalar problem with source terms. In Section 3, we present the main algorithm of the parametrized PP finite difference WENO RK method for the compressible Euler equation in one and two dimensions. An implementation procedure is given in the presence of source terms. In Section 4, we perform extensive numerical tests to illustrate the effectiveness of the proposed method. We finally conclude in Section 5.

2 Parametrized MPP flux limiters for scalar equations

2.1 Review of MPP flux limiters for scalar equations

For simplicity, we consider a simple one-dimensional hyperbolic conservation equation

$$u_t + f(u)_x = 0, \quad x \in [0, 1], \quad (2.1)$$

with an initial condition $u(x, 0) = u_0(x)$ and a periodic boundary condition. We adopt the following spatial discretization for the domain $[0, 1]$

$$0 = x_{\frac{1}{2}} < x_{\frac{3}{2}} < \cdots < x_{N+\frac{1}{2}} = 1,$$

where $I_j = [x_{j-\frac{1}{2}}, x_{j+\frac{1}{2}}]$ has the mesh size $\Delta x = \frac{1}{N}$. Let $u_j(t)$ denote the solution at grid point $x_j = \frac{1}{2}(x_{j-\frac{1}{2}} + x_{j+\frac{1}{2}})$ at continuous time t . The finite difference scheme evolves the point values of the solution in a conservative form

$$\frac{d}{dt}u_j(t) + \frac{1}{\Delta x}(\hat{H}_{j+1/2} - \hat{H}_{j-1/2}) = 0. \quad (2.2)$$

The numerical flux $\hat{H}_{j+\frac{1}{2}}$ in equation (2.2) can be reconstructed from neighboring flux functions $f(u(x_i, t))$, $i = j - p, \dots, j + q$ with high order by WENO reconstructions [9, 14]. By adaptively assigning nonlinear weights to neighboring candidate stencils, the WENO reconstruction preserves high order accuracy of the linear scheme around smooth regions of the solution, while producing a sharp and essentially non-oscillatory capture of discontinuities. Equation (2.2) can be further discretized in time by a high order time integrator via the method-of-line approach. For example, the scheme with a third order total variation diminishing (TVD) RK time discretization is

$$\begin{aligned} u_j^{(1)} &= u_j^n + \Delta t L(u_j^n), \\ u_j^{(2)} &= u_j^n + \frac{1}{4}\Delta t (L(u_j^n) + L(u_j^{(1)})), \\ u_j^{n+1} &= u_j^n + \frac{1}{6}\Delta t (L(u_j^n) + L(u_j^{(1)}) + 4L(u_j^{(2)})). \end{aligned} \quad (2.3)$$

where $u_j^{(k)}$ and u_j^n denotes the numerical solution at x_j at k^{th} RK stage and at time t^n respectively. Let Δt be the time step size. $L(u^{(k)}) \doteq -\frac{1}{\Delta x}(\hat{H}_{j+\frac{1}{2}}^{(k)} - \hat{H}_{j-\frac{1}{2}}^{(k)})$ with $\hat{H}_{j+\frac{1}{2}}^{(k)}$ being the numerical flux from finite difference WENO reconstruction based on $\{u_j^{(k)}\}_{j=1}^N$ at intermedia RK stages. Equation (2.3) in the final stage of RK method can be re-written as

$$u_j^{n+1} = u_j^n - \lambda(\hat{H}_{j+\frac{1}{2}}^{rk} - \hat{H}_{j-\frac{1}{2}}^{rk}), \quad (2.4)$$

with $\lambda = \frac{\Delta t}{\Delta x}$ and

$$\hat{H}_{j+\frac{1}{2}}^{rk} \doteq \frac{1}{6} \left(\hat{H}_{j+\frac{1}{2}}^n + \hat{H}_{j+\frac{1}{2}}^{(1)} + 4\hat{H}_{j+\frac{1}{2}}^{(2)} \right). \quad (2.5)$$

The parametrized MPP flux limiters in [20] is based on the finite difference RK WENO scheme for equation (2.1) reviewed above. Let $u_m = \min_x(u(x, 0))$ and $u_M = \max_x(u(x, 0))$. The idea of the parametrized MPP flux limiter is to modify the high order flux $\hat{H}_{j+\frac{1}{2}}^{rk}$ in equation (2.5) towards a first order monotone flux denoted as $\hat{h}_{j+\frac{1}{2}}$ by taking a linear combination of them,

$$\tilde{H}_{j+\frac{1}{2}}^{rk} \doteq \hat{h}_{j+\frac{1}{2}} + \theta_{j+\frac{1}{2}}(\hat{H}_{j+\frac{1}{2}}^{rk} - \hat{h}_{j+\frac{1}{2}}), \quad \theta_{j+\frac{1}{2}} \in [0, 1]. \quad (2.6)$$

the original high order flux $\hat{H}_{j+\frac{1}{2}}^{rk}$ in equation (2.5) is then replaced by the modified flux $\tilde{H}_{j+\frac{1}{2}}^{rk}$ above.

To preserve the MPP property, we wish to have $u_m \leq u_j^{n+1} \leq u_M$ at the final RK stage on each time step, i.e.

$$u_m \leq u_j^n - \lambda(\tilde{H}_{j+\frac{1}{2}}^{rk} - \tilde{H}_{j-\frac{1}{2}}^{rk}) \leq u_M. \quad (2.7)$$

For the parametrized MPP flux limiter, a pair $(\Lambda_{-\frac{1}{2}, I_j}, \Lambda_{+\frac{1}{2}, I_j})$ needs to be found such that any pair $(\theta_{j-\frac{1}{2}}, \theta_{j+\frac{1}{2}}) \in [0, \Lambda_{-\frac{1}{2}, I_j}] \times [0, \Lambda_{+\frac{1}{2}, I_j}]$ satisfies (2.7). Under such a constraint, $\theta_{j\pm\frac{1}{2}}$ are chosen to be as close to 1 as possible for accuracy, which is done by the following three steps. Below ϵ is a small positive number to avoid the denominator to be 0, e.g., $\epsilon = 10^{-13}$.

1. The right inequality of (2.7), that is the maximum value part, can be rewritten as

$$\lambda\theta_{j-\frac{1}{2}}(\hat{H}_{j-\frac{1}{2}}^{rk} - \hat{h}_{j-\frac{1}{2}}) - \lambda\theta_{j+\frac{1}{2}}(\hat{H}_{j+\frac{1}{2}}^{rk} - \hat{h}_{j+\frac{1}{2}}) - \Gamma_j^M \leq 0, \quad (2.8)$$

where $\Gamma_j^M = u_M - u_j + \lambda(\hat{h}_{j+\frac{1}{2}} - \hat{h}_{j-\frac{1}{2}}) \geq 0$. Let $F_{j-\frac{1}{2}} = \hat{H}_{j-\frac{1}{2}}^{rk} - \hat{h}_{j-\frac{1}{2}}$, the decoupling of (2.8) on cell I_j gives:

- (a) If $F_{j-\frac{1}{2}} \leq 0$ and $F_{j+\frac{1}{2}} \geq 0$, let $(\Lambda_{-\frac{1}{2}, I_j}^M, \Lambda_{+\frac{1}{2}, I_j}^M) = (1, 1)$.
- (b) If $F_{j-\frac{1}{2}} \leq 0$ and $F_{j+\frac{1}{2}} < 0$, let $(\Lambda_{-\frac{1}{2}, I_j}^M, \Lambda_{+\frac{1}{2}, I_j}^M) = (1, \min(1, \frac{\Gamma_j^M}{-\lambda F_{j+\frac{1}{2}} + \epsilon}))$.
- (c) If $F_{j-\frac{1}{2}} > 0$ and $F_{j+\frac{1}{2}} \geq 0$, let $(\Lambda_{-\frac{1}{2}, I_j}^M, \Lambda_{+\frac{1}{2}, I_j}^M) = (\min(1, \frac{\Gamma_j^M}{\lambda F_{j-\frac{1}{2}} + \epsilon}), 1)$.
- (d) If $F_{j-\frac{1}{2}} > 0$ and $F_{j+\frac{1}{2}} < 0$,
 - if $(\theta_{j-\frac{1}{2}}, \theta_{j+\frac{1}{2}}) = (1, 1)$ satisfies (2.8), let $(\Lambda_{-\frac{1}{2}, I_j}^M, \Lambda_{+\frac{1}{2}, I_j}^M) = (1, 1)$;
 - otherwise, let $(\Lambda_{-\frac{1}{2}, I_j}^M, \Lambda_{+\frac{1}{2}, I_j}^M) = (\frac{\Gamma_j^M}{\lambda F_{j-\frac{1}{2}} - \lambda F_{j+\frac{1}{2}} + \epsilon}, \frac{\Gamma_j^M}{\lambda F_{j-\frac{1}{2}} - \lambda F_{j+\frac{1}{2}} + \epsilon})$.

2. The left inequality of (2.7), that is the minimum value part, can be rewritten as

$$0 \leq \lambda \theta_{j-\frac{1}{2}} (\hat{H}_{j-\frac{1}{2}}^{rk} - \hat{h}_{j-\frac{1}{2}}) - \lambda \theta_{j+\frac{1}{2}} (\hat{H}_{j+\frac{1}{2}}^{rk} - \hat{h}_{j+\frac{1}{2}}) - \Gamma_j^m, \quad (2.9)$$

where $\Gamma_j^m = u_m - u_j + \lambda(\hat{h}_{j+\frac{1}{2}} - \hat{h}_{j-\frac{1}{2}}) \leq 0$. Similar to the maximum value case, the decoupling of (2.9) on cell I_j gives:

- (a) If $F_{j-\frac{1}{2}} \geq 0$ and $F_{j+\frac{1}{2}} \leq 0$, let $(\Lambda_{-\frac{1}{2}, I_j}^m, \Lambda_{+\frac{1}{2}, I_j}^m) = (1, 1)$;
- (b) If $F_{j-\frac{1}{2}} \geq 0$ and $F_{j+\frac{1}{2}} > 0$, let $(\Lambda_{-\frac{1}{2}, I_j}^m, \Lambda_{+\frac{1}{2}, I_j}^m) = (1, \min(1, \frac{\Gamma_j^m}{-\lambda F_{j+\frac{1}{2}} - \epsilon}))$;
- (c) If $F_{j-\frac{1}{2}} < 0$ and $F_{j+\frac{1}{2}} \leq 0$, let $(\Lambda_{-\frac{1}{2}, I_j}^m, \Lambda_{+\frac{1}{2}, I_j}^m) = (\min(1, \frac{\Gamma_j^m}{\lambda F_{j-\frac{1}{2}} - \epsilon}), 1)$;
- (d) If $F_{j-\frac{1}{2}} < 0$ and $F_{j+\frac{1}{2}} > 0$,
 - when $(\theta_{j-\frac{1}{2}}, \theta_{j+\frac{1}{2}}) = (1, 1)$ satisfies (2.9), let $(\Lambda_{-\frac{1}{2}, I_j}^m, \Lambda_{+\frac{1}{2}, I_j}^m) = (1, 1)$;
 - otherwise, let $(\Lambda_{-\frac{1}{2}, I_j}^m, \Lambda_{+\frac{1}{2}, I_j}^m) = (\frac{\Gamma_j^m}{\lambda F_{j-\frac{1}{2}} - \lambda F_{j+\frac{1}{2}} - \epsilon}, \frac{\Gamma_j^m}{\lambda F_{j-\frac{1}{2}} - \lambda F_{j+\frac{1}{2}} - \epsilon})$.

3. The locally defined limiting parameter is given as

$$\Lambda_{j+\frac{1}{2}} = \min(\Lambda_{+\frac{1}{2}, I_j}^M, \Lambda_{-\frac{1}{2}, I_{j+1}}^M, \Lambda_{+\frac{1}{2}, I_j}^m, \Lambda_{-\frac{1}{2}, I_{j+1}}^m), \quad j = 0, \dots, N. \quad (2.10)$$

The flux limiting procedure above guarantees the MPP property of the numerical solution by the design. It is theoretically proved to preserve up to fourth order spatial and temporal accuracy for smooth solutions [20, 21].

2.2 Scalar advection equations with source terms

We consider scalar advection problems with a source term

$$u_t + f(u)_x = s(u). \quad (2.11)$$

In particular, we consider the class of problems whose solutions enjoy the PP property, that is, the lower bound of the solution is 0 (such kind of problem might not preserve the MPP property). For example, when $s(u) = -ku$ with a positive k , with positive initial values and periodic boundary conditions, the solution satisfies the PP property. The flux limiter is designed base on the PP property of a first order scheme

$$u_j^{n+1} = u_j^n - \lambda(\hat{h}_{j+\frac{1}{2}} - \hat{h}_{j-\frac{1}{2}}) + \Delta t s(u_j^n), \quad (2.12)$$

under the time step constraint

$$\Delta t \leq \frac{\text{CFL } \Delta x}{\lambda_{\max} + s_{\max} \Delta x}, \quad (2.13)$$

where $\lambda_{max} = \max |f'(u)|$ and $s_{max} = \max |s'(u)|$.

We propose to first modify the source term such that $\tilde{u}_j^{n+1} \geq \epsilon_s$, with $\epsilon_s = \min_j(u_j^{n+1}, 10^{-13})$, where $\{u_j^{n+1}\}$ are positive solutions computed from (2.12) and 10^{-13} is a small positive number related to machine precision. Here \tilde{u}_j^{n+1} is

$$\tilde{u}_j^{n+1} = u_j^n - \lambda(\hat{h}_{j+\frac{1}{2}} - \hat{h}_{j-\frac{1}{2}}) + \Delta t \tilde{s}_j^{rk}, \quad (2.14)$$

with

$$\tilde{s}_j^{rk} = r_j(\hat{s}_j^{rk} - s(u_j^n)) + s(u_j^n), \quad (2.15)$$

and

$$\hat{s}_j^{rk} \doteq \frac{1}{6} \left(s(u_j^n) + s(u_j^{(1)}) + 4s(u_j^{(2)}) \right), \quad (2.16)$$

as in (2.5). r_j is designed by the linear constraints to preserve the PP property of $\{\tilde{u}_j^{n+1}\}_j$. Specifically,

$$r_j = \begin{cases} \min(\frac{\epsilon_s - u_j^{n+1}}{\Delta t \Delta s_j}, 1), & \text{if } \tilde{u}_j < \epsilon_s \\ 1, & \text{otherwise} \end{cases},$$

where $\Delta s_j = \hat{s}_j^{rk} - s(u_j^n)$ and $\tilde{u}_j = u_j^n - \lambda(\hat{h}_{j+\frac{1}{2}} - \hat{h}_{j-\frac{1}{2}}) + \Delta t \hat{s}_j^{rk}$. Next the parametrized MPP flux limiters are applied as in (2.7) to satisfy

$$\epsilon_s \leq u_j^n - \lambda(\tilde{H}_{j+\frac{1}{2}}^{rk} - \tilde{H}_{j-\frac{1}{2}}^{rk}) + \Delta t \tilde{s}_j^{rk}. \quad (2.17)$$

(2.17) leads to the same decomposed inequality (2.9) for the minimum value part, only with Γ_j^m given by

$$\Gamma_j^m = \epsilon_s - u_j + \lambda(\hat{h}_{j+\frac{1}{2}} - \hat{h}_{j-\frac{1}{2}}) - \Delta t \tilde{s}_j^{rk} \leq 0. \quad (2.18)$$

The procedure proposed above for treating equations with a source term is PP by the design, and is shown to maintain high order accuracy by numerical tests in Section 4.

3 Parametrized PP flux limiters for compressible Euler equations

In this section, we first extend the parametrized MPP flux limiters to PP flux limiters for the compressible Euler equations. We then describe how to generalize the proposed approach to systems with source terms and to high dimensional systems. In this section, we use letters in bold for vectors.

3.1 Parametrized positivity preserving flux limiters

For compressible Euler equations in one dimension

$$\mathbf{u}_t + \mathbf{f}(\mathbf{u})_x = 0, \quad (3.1)$$

with $\mathbf{u} = (\rho, \rho u, E)^T$, $\mathbf{f}(\mathbf{u}) = (\rho u, \rho u^2 + p, (E + p)u)^T$, where ρ is the density, u is the velocity, p is the pressure, $m = \rho u$ is the momentum, $E = \frac{1}{2}\rho u^2 + \frac{p}{\gamma-1}$ is the total energy from equation of state (EOS) and γ is the ratio of specific heat ($\gamma = 1.4$ for the air). Denote $\hat{\mathbf{h}}_{j+\frac{1}{2}}$ to be a first order monotone flux, and $\hat{\mathbf{H}}_{j+\frac{1}{2}}^{rk}$ to be the linear combinations of fluxes from multiple RK stages, similar to equation (2.5), but in a component-by-component fashion. For positivity preserving, we are seeking the flux limiters of the type

$$\tilde{\mathbf{H}}_{j+\frac{1}{2}}^{rk} = \theta_{j+\frac{1}{2}}(\hat{\mathbf{H}}_{j+\frac{1}{2}}^{rk} - \hat{\mathbf{h}}_{j+\frac{1}{2}}) + \hat{\mathbf{h}}_{j+\frac{1}{2}} \quad (3.2)$$

such that

$$\begin{cases} \rho_j^{n+1} > 0, \\ p_j^{n+1} > 0, \end{cases} \quad (3.3)$$

for the updated solution

$$\mathbf{u}_j^{n+1} = \mathbf{u}_j^n - \lambda(\tilde{\mathbf{H}}_{j+\frac{1}{2}}^{rk} - \tilde{\mathbf{H}}_{j-\frac{1}{2}}^{rk}). \quad (3.4)$$

In the parametrized flux limiters' framework, a pair of $(\Lambda_{-\frac{1}{2}, I_j}, \Lambda_{+\frac{1}{2}, I_j})$ is found such that the updated solution satisfies (3.3) for any $(\theta_{j-\frac{1}{2}}, \theta_{j+\frac{1}{2}}) \in [0, \Lambda_{-\frac{1}{2}, I_j}] \times [0, \Lambda_{+\frac{1}{2}, I_j}]$. The high order flux $\hat{\mathbf{H}}_{j+\frac{1}{2}}^{rk}$ is modified by (3.2) to preserve positive density and pressure. In simulations, preserving positivity is implemented by

$$\begin{cases} \rho_j^{n+1} \geq \epsilon_\rho, \\ p_j^{n+1} \geq \epsilon_p. \end{cases} \quad (3.5)$$

where we introduce small positive numbers ϵ_ρ defined by $\min_j(\rho_j^{n+1}, 10^{-13})$ and ϵ_p defined by $\min_j(p_j^{n+1}, 10^{-13})$. ρ_j^{n+1} and p_j^{n+1} are positive density and pressure obtained by the first order monotone scheme and 10^{-13} is related to the machine precision. Let us denote the first order monotone flux by $\hat{\mathbf{h}}(\mathbf{u}) = (f^\rho, f^m, f^E)^T$, similarly $\hat{\mathbf{H}}^{rk} = (\hat{f}^\rho, \hat{f}^m, \hat{f}^E)^T$ and $\tilde{\mathbf{H}}^{rk} = (\tilde{f}^\rho, \tilde{f}^m, \tilde{f}^E)^T$. The proposed process can be dissected into two steps.

1. Find the limiting parameters $\theta_{j\pm\frac{1}{2}}$ to preserve the positivity of the density,

$$\rho_j^{n+1} = \rho_j^n - \lambda(\tilde{f}_{j+\frac{1}{2}}^\rho - \tilde{f}_{j-\frac{1}{2}}^\rho). \quad (3.6)$$

Thus, the limiting parameters $\theta_{j\pm\frac{1}{2}}$ are found to satisfy

$$\epsilon_\rho \leq \Gamma_j - \lambda(\theta_{j+\frac{1}{2}}(\hat{f}_{j+\frac{1}{2}}^\rho - f_{j+\frac{1}{2}}^\rho) - \theta_{j-\frac{1}{2}}(\hat{f}_{j-\frac{1}{2}}^\rho - f_{j-\frac{1}{2}}^\rho)), \quad (3.7)$$

which is equivalent to

$$0 \leq \Gamma_j - \epsilon_\rho - \lambda(\theta_{j+\frac{1}{2}}(\hat{f}_{j+\frac{1}{2}}^\rho - f_{j+\frac{1}{2}}^\rho) - \theta_{j-\frac{1}{2}}(\hat{f}_{j-\frac{1}{2}}^\rho - f_{j-\frac{1}{2}}^\rho)), \quad (3.8)$$

where $\Gamma_j = \rho_j^n - \lambda(f_{j+\frac{1}{2}}^\rho - f_{j-\frac{1}{2}}^\rho) \geq \epsilon_\rho$. A pair of limiting parameters $(\Lambda_{-\frac{1}{2},I_j}^\rho, \Lambda_{+\frac{1}{2},I_j}^\rho)$ for the positive density of (3.8) can be identified by a similar procedure as described in Section 2.1. We can define a set for the positive density ρ_j^{n+1}

$$S_\rho = \{(\theta_{j-\frac{1}{2}}, \theta_{j+\frac{1}{2}}) : 0 \leq \theta_{j-\frac{1}{2}} \leq \Lambda_{-\frac{1}{2},I_j}^\rho, 0 \leq \theta_{j+\frac{1}{2}} \leq \Lambda_{+\frac{1}{2},I_j}^\rho\}, \quad (3.9)$$

which is plotted as the rectangle bounded by the dash line in Figure 3.1.

2. Find the limiting parameters $\theta_{j\pm\frac{1}{2}}$ within the region S_ρ to preserve the positivity of the pressure. We seek a sufficient condition such that the pressure given by (3.4) satisfies

$$p_j^{n+1}(\theta_{j-\frac{1}{2}}, \theta_{j+\frac{1}{2}}) = (\gamma - 1) \left(E_j^{n+1} - \frac{1}{2} \frac{(m_j^{n+1})^2}{\rho_j^{n+1}} \right) \geq \epsilon_p. \quad (3.10)$$

The decoupling of (3.10) for $(\theta_{j-\frac{1}{2}}, \theta_{j+\frac{1}{2}})$ is different from the scalar case since the principal variables are nonlinearly dependent on each other. However the idea is still to separate $\theta_{j-\frac{1}{2}}$ and $\theta_{j+\frac{1}{2}}$. Since $\rho_j^{n+1} \geq \epsilon_\rho$ is guaranteed by the previous step, we first put the concave property of pressure [23] in the following remark for future reference:

Remark 3.1. The pressure as a function of (ρ, m, E) is concave, i.e., $p(\alpha \mathbf{U}_1 + (1 - \alpha) \mathbf{U}_2) \geq \alpha p(\mathbf{U}_1) + (1 - \alpha)p(\mathbf{U}_2)$ for $0 \leq \alpha \leq 1$ if $\rho_1, \rho_2 > 0$. Therefore $p_j^{n+1}(\theta_{j-\frac{1}{2}}, \theta_{j+\frac{1}{2}})$ is a concave function of $(\theta_{j-\frac{1}{2}}, \theta_{j+\frac{1}{2}})$ on S_ρ due to the linear dependence of $(\rho_j^{n+1}, m_j^{n+1}, E_j^{n+1})$ on $(\theta_{j-\frac{1}{2}}, \theta_{j+\frac{1}{2}})$. Therefore, if $p_j^{n+1}(\vec{\theta}^l) \geq \epsilon_p$, with $\vec{\theta}^l = (\theta_{j-\frac{1}{2}}^l, \theta_{j+\frac{1}{2}}^l)$ for $l = 1, 2$, then $p_j^{n+1}(\alpha \vec{\theta}^1 + (1 - \alpha) \vec{\theta}^2) \geq \epsilon_p$, $0 \leq \alpha \leq 1$.

We define an admissible set

$$S_\theta = \{(\theta_{j-\frac{1}{2}}, \theta_{j+\frac{1}{2}}) \in S_\rho : (\theta_{j-\frac{1}{2}}, \theta_{j+\frac{1}{2}}) \text{ satisfies (3.10)}\}. \quad (3.11)$$

S_θ is a convex set thanks to Remark 3.1. Let the three vertices of the rectangle S_ρ other than $(0, 0)$ be denoted by

$$A^1 = (0, \Lambda_{+\frac{1}{2},I_j}^\rho), \quad A^2 = (\Lambda_{-\frac{1}{2},I_j}^\rho, 0), \quad A^3 = (\Lambda_{-\frac{1}{2},I_j}^\rho, \Lambda_{+\frac{1}{2},I_j}^\rho), \quad (3.12)$$

see Figure 3.1. Based on the concave property in Remark 3.1, we propose the following way of decoupling (3.10).

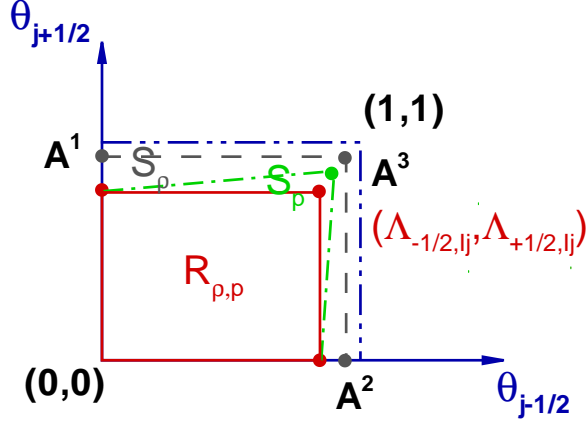


Figure 3.1: The decoupling rectangle $R_{\rho,p}$ (bounded by the solid line) with the right-top node $(\Lambda_{-\frac{1}{2},I_j}, \Lambda_{+\frac{1}{2},I_j})$. S_ρ is the rectangle bounded by the dash line. S_p is the polygonal bounded by the dash dot line.

- (a) For $i=1, 2, 3$, if $p(A^i) \geq \epsilon_p$, let $B^i = A^i$; otherwise find r such that $p(rA^i) \geq \epsilon_p$ and let $B^i = rA^i$. The three B^i 's and $(0,0)$ form a convex polygonal region, denoted as S_p , inside S_θ . Such convex polygonal region S_p is outlined by the dash dot line in Figure 3.1.
- (b) We define the decoupling rectangle, as a subset of S_p , to be

$$R_{\rho,p} = [0, \min(B_1^2, B_1^3)] \times [0, \min(B_2^1, B_2^3)], \quad (3.13)$$

see the region outlined by the solid line in Figure 3.1. That is, within S_p , we find the decoupling rectangle $R_{\rho,p}$ with left-bottom node on $(0,0)$ and right-top node $(\Lambda_{-\frac{1}{2},I_j}, \Lambda_{+\frac{1}{2},I_j})$ as close to $(1,1)$ as possible to best preserve the accuracy while achieving the PP property of high order numerical schemes. Let

$$(\Lambda_{-\frac{1}{2},I_j}, \Lambda_{+\frac{1}{2},I_j}) = (\min(B_1^2, B_1^3), \min(B_2^1, B_2^3)). \quad (3.14)$$

Finally, similar to equation (2.10) for the MPP flux limiters, the locally defined limiting parameter is given as $\theta_{j+\frac{1}{2}} = \min(\Lambda_{-\frac{1}{2},I_j}, \Lambda_{+\frac{1}{2},I_{j+1}})$.

Remark 3.2. The limiter above can preserve positive density and pressure by its design due to the two sufficient conditions (3.7) and (3.10). For general equation of state, if $\rho > 0$, then $p > 0 \Leftrightarrow e > 0$, where the internal energy e can always be written as a concave function of $(\rho, m, E)^T$ similarly as (3.10) [26]. Similar procedure can be followed for PP property of numerical solutions.

3.2 Extension to Euler system with source term

The compressible Euler equations may come with source terms in the form of

$$\mathbf{u}_t + \mathbf{f}(\mathbf{u})_x = \mathbf{s}(\mathbf{u}), \quad (3.15)$$

For example, four kinds of source terms were discussed in [25]: geometric, gravity, chemical reaction and radiative cooling. The PP flux limiters can be applied by the following three steps.

1. Choose a time step, such that the first order scheme (3.16) is PP,

$$\mathbf{u}_j^{n+1} = \mathbf{u}_j^n - \lambda(\hat{\mathbf{h}}_{j+\frac{1}{2}} - \hat{\mathbf{h}}_{j-\frac{1}{2}}) + \Delta t \mathbf{s}(\mathbf{u}_j^n). \quad (3.16)$$

2. Find r such that the scheme (3.17) with the modified source terms is PP

$$\mathbf{u}_j^{n+1} = \mathbf{u}_j^n - \lambda(\hat{\mathbf{h}}_{j+\frac{1}{2}} - \hat{\mathbf{h}}_{j-\frac{1}{2}}) + \Delta t \tilde{\mathbf{s}}_j^{rk}, \quad (3.17)$$

with $\tilde{\mathbf{s}}_j^{rk} = r(\hat{\mathbf{s}}_j^{rk} - \mathbf{s}(\mathbf{u}_j^n)) + \mathbf{s}(\mathbf{u}_j^n)$, $\hat{\mathbf{s}}_j^{rk}$ is similarly defined as (2.16) component-by-component.

3. Finally find $\theta_{j\pm\frac{1}{2}}$ for the modified high order flux $\tilde{\mathbf{H}}_{j\pm\frac{1}{2}}^{rk}$, such that (3.18) is PP

$$\mathbf{u}_j^{n+1} = \mathbf{u}_j^n - \lambda(\tilde{\mathbf{H}}_{j+\frac{1}{2}}^{rk} - \tilde{\mathbf{H}}_{j-\frac{1}{2}}^{rk}) + \Delta t \tilde{\mathbf{s}}_j^{rk}. \quad (3.18)$$

The procedure is similar as in the previous subsection.

3.3 Extension to the multi-dimensional Euler system

In this subsection, we extend the previously proposed PP flux limiters to Euler equations in two-dimensions

$$\mathbf{u}_t + \mathbf{f}(\mathbf{u})_x + \mathbf{g}(\mathbf{u})_y = 0, \quad (3.19)$$

with $\mathbf{u} = (\rho, m_u, m_v, E)^T$, $\mathbf{f}(\mathbf{u}) = (m_u, \rho u^2 + p, \rho uv, (E + p)u)^T$ and $\mathbf{g}(\mathbf{u}) = (m_v, \rho uv, \rho v^2 + p, (E + p)v)^T$. ρ is the density, u is the velocity in x direction, v is the velocity in y direction,

p is the pressure, $m_u = \rho u$ and $m_v = \rho v$ are the momenta, $E = \frac{1}{2}\rho u^2 + \frac{1}{2}\rho v^2 + \frac{p}{\gamma-1}$ is the total energy and γ is the ratio of specific heat.

The high order finite difference scheme with PP flux limiters at the final stage of a RK time discretization is given by

$$\mathbf{u}_{i,j}^{n+1} = \mathbf{u}_{i,j}^n - \lambda_x(\tilde{\mathbf{H}}_{i+\frac{1}{2},j}^{rk} - \tilde{\mathbf{H}}_{i-\frac{1}{2},j}^{rk}) - \lambda_y(\tilde{\mathbf{G}}_{i,j+\frac{1}{2}}^{rk} - \tilde{\mathbf{G}}_{i,j-\frac{1}{2}}^{rk}), \quad (3.20)$$

with

$$\tilde{\mathbf{H}}_{i+\frac{1}{2},j}^{rk} = \theta_{i+\frac{1}{2},j}(\hat{\mathbf{H}}_{i+\frac{1}{2},j}^{rk} - \hat{\mathbf{h}}_{i+\frac{1}{2},j}) + \hat{\mathbf{h}}_{i+\frac{1}{2},j}, \quad (3.21)$$

$$\tilde{\mathbf{G}}_{i,j+\frac{1}{2}}^{rk} = \theta_{i,j+\frac{1}{2}}(\hat{\mathbf{G}}_{i,j+\frac{1}{2}}^{rk} - \hat{\mathbf{g}}_{i,j+\frac{1}{2}}) + \hat{\mathbf{g}}_{i,j+\frac{1}{2}}, \quad (3.22)$$

where $\hat{\mathbf{H}}_{i+\frac{1}{2},j}^{rk}$ and $\hat{\mathbf{G}}_{i,j+\frac{1}{2}}^{rk}$ are linear combinations of fluxes from multiple RK stages similarly as (2.5) in the scalar case but in a component-wise fashion, $\hat{\mathbf{h}}_{i+\frac{1}{2},j}$ and $\hat{\mathbf{g}}_{i,j+\frac{1}{2}}$ are first order monotone fluxes.

Similar to the 1D case, we find the four parametrized limiters $\Lambda_{L,Iij}^\rho$, $\Lambda_{R,Iij}^\rho$, $\Lambda_{U,Iij}^\rho$ and $\Lambda_{D,Iij}^\rho$, such that for all $\theta_{i\pm\frac{1}{2},j}$ and $\theta_{i,j\pm\frac{1}{2}}$ in the set

$$S_\rho = \{(\theta_{i-\frac{1}{2},j}, \theta_{i+\frac{1}{2},j}, \theta_{i,j-\frac{1}{2}}, \theta_{i,j+\frac{1}{2}}) : 0 \leq \theta_{i-\frac{1}{2},j} \leq \Lambda_{L,Iij}^\rho, \\ 0 \leq \theta_{i+\frac{1}{2},j} \leq \Lambda_{R,Iij}^\rho, 0 \leq \theta_{i,j-\frac{1}{2}} \leq \Lambda_{D,Iij}^\rho, 0 \leq \theta_{i,j+\frac{1}{2}} \leq \Lambda_{U,Iij}^\rho\} \quad (3.23)$$

we have $\rho_{i,j}^{n+1} \geq \epsilon_\rho$. With the positive density $\rho_{i,j}^{n+1}$, the pressure is updated by the constraint

$$p_{i,j}^{n+1}(\theta_{i-\frac{1}{2},j}, \theta_{i+\frac{1}{2},j}, \theta_{i,j-\frac{1}{2}}, \theta_{i,j+\frac{1}{2}}) = \\ (\gamma - 1)(E_{i,j}^{n+1} - \frac{1}{2} \frac{((m_u)_{i,j}^{n+1})^2 + ((m_v)_{i,j}^{n+1})^2}{\rho_{i,j}^{n+1}}) \geq \epsilon_p. \quad (3.24)$$

Let the convex admissible set for positive pressure be

$$S_\theta = \{(\theta_{i-\frac{1}{2},j}, \theta_{i+\frac{1}{2},j}, \theta_{i,j-\frac{1}{2}}, \theta_{i,j+\frac{1}{2}}) \in S_\rho : (\theta_{i-\frac{1}{2},j}, \theta_{i+\frac{1}{2},j}, \theta_{i,j-\frac{1}{2}}, \theta_{i,j+\frac{1}{2}}) \text{ satisfies (3.24)}\} \quad (3.25)$$

Let the sixteen vertices of S_ρ denoted by

$$A^{k_1,k_2,k_3,k_4} = (k_1\Lambda_{L,Iij}^\rho, k_2\Lambda_{R,Iij}^\rho, k_3\Lambda_{D,Iij}^\rho, k_4\Lambda_{U,Iij}^\rho), \quad (3.26)$$

with k_1, k_2, k_3, k_4 to be 0 or 1. We decouple (3.24) in the following way:

1. For $(k_1, k_2, k_3, k_4) \neq (0, 0, 0, 0)$, if $p(A^{k_1,k_2,k_3,k_4}) \geq \epsilon_p$, let $B^{k_1,k_2,k_3,k_4} = A^{k_1,k_2,k_3,k_4}$; otherwise find r such that $P(rA^{k_1,k_2,k_3,k_4}) \geq \epsilon_p$ and let $B^{k_1,k_2,k_3,k_4} = rA^{k_1,k_2,k_3,k_4}$. The 15 B^{k_1,k_2,k_3,k_4} 's with the origin $(0, 0, 0, 0)$ form a four dimensional polyhedra inside S_θ ;

2. The decoupling tesseract can be defined by

$$R_{\rho,p} = [0, \min(B_1^{1,1,1,0}, B_1^{1,1,0,1}, B_1^{1,0,1,1})] \times [0, \min(B_2^{1,1,1,0}, B_2^{1,1,0,1}, B_2^{0,1,1,1})] \\ \times [0, \min(B_3^{1,1,1,0}, B_3^{1,0,1,1}, B_3^{0,1,1,1})] \times [0, \min(B_4^{1,1,0,1}, B_4^{1,0,1,1}, B_4^{0,1,1,1})]. \quad (3.27)$$

Let

$$(\Lambda_{L,I_{ij}}, \Lambda_{R,I_{ij}}, \Lambda_{D,I_{ij}}, \Lambda_{U,I_{ij}}) = (\min(B_1^{1,1,1,0}, B_1^{1,1,0,1}, B_1^{1,0,1,1}), \min(B_2^{1,1,1,0}, B_2^{1,1,0,1}, B_2^{0,1,1,1}), \min(B_3^{1,1,1,0}, B_3^{1,0,1,1}, B_3^{0,1,1,1}), \min(B_4^{1,1,0,1}, B_4^{1,0,1,1}, B_4^{0,1,1,1})). \quad (3.28)$$

Finally, similar to equation (2.10) for the MPP flux limiters, the locally defined limiting parameter is given as $\theta_{i+\frac{1}{2},j} = \min(\Lambda_{L,I_{ij}}, \Lambda_{R,I_{i+1,j}})$ and $\theta_{i,j+\frac{1}{2}} = \min(\Lambda_{D,I_{ij}}, \Lambda_{U,I_{i,j+1}})$.

Remark 3.3. For two dimensional compressible Euler equations with source terms, it can be done similarly as the one dimensional case.

4 Numerical simulations

In this section, we will use the 5th order finite difference WENO scheme for space discretization [9] and a 4th order Runge-Kutta time discretization [15], denote as “WENO5RK4”, with the proposed PP flux limiters for simulating the compressible Euler equations. Here a 4th order RK time discretization is adopted for better observation of accuracy by taking the time step to be $\Delta t = \text{CFL} \Delta x$. Most of the tests are from [26]. Below, $\text{CFL} = 0.6$ unless otherwise specified.

Example 4.1. (Accuracy test for a scalar problem with a source term.) We consider $u_t + u_x = -u$ with the initial condition

$$u(x, 0) = \sin^4(x),$$

and the periodic boundary condition. The exact solution is given by

$$u(x, t) = e^{-t} \sin^4(x - t).$$

The minimum value of the exact solution is $u_m = 0$. This example is used to test the PP property and accuracy of dealing with a source term. In Table 4.1, we can see the PP property is preserved and the 5th order accuracy has been maintained.

Table 4.1: Example 4.1. A scalar advection problem with a source term at $T = 0.1$. v_{min} is the minimum value of the numerical solution.

	N	L^1 error	order	L^∞ error	order	v_{min}
without limiters	40	3.36E-04	–	8.78E-04	–	-1.35E-04
	80	2.03E-05	4.05	1.24E-04	2.82	-1.05E-05
	160	6.75E-07	4.91	6.25E-06	4.31	-1.88E-06
	320	1.67E-08	5.34	1.29E-07	5.60	-3.02E-09
	640	4.30E-10	5.28	2.60E-09	5.63	-5.28E-11
with limiters	40	3.25E-04	–	8.66E-04	–	5.67E-15
	80	1.92E-05	4.08	1.17E-04	2.89	1.18E-05
	160	6.38E-07	4.91	6.25E-06	4.22	3.01E-16
	320	1.67E-08	5.26	1.29E-07	5.60	6.33E-10
	640	4.31E-10	5.28	2.60E-09	5.63	3.46E-11

Example 4.2. (Accuracy test for the global Lax-Friedrichs flux.) We consider the Burgers' equation with the initial condition

$$u(x, 0) = (1 + \sin(x))/2$$

and a periodic boundary condition. We consider the WENO5RK4 scheme with the global Lax-Friedrichs (LxF) fluxes. Let

$$f_i^\pm = \frac{1}{2}(f(u_i^n) \pm \alpha u_i^n), \quad i = j - p, \dots, j + q, \quad (4.1)$$

with $\alpha \geq \max_u |f'(u)|$. The numerical flux $\hat{H}_{j+\frac{1}{2}} = f_{j+\frac{1}{2}}^- + f_{j+\frac{1}{2}}^+$ in (2.2), where $f_{j+\frac{1}{2}}^\pm$ are reconstructed based on WENO schemes from (4.1) with the corresponding upwind mechanism. We numerically investigate the time step restriction for maintaining high order accuracy using the global Lax-Friedrichs flux, since it is frequently used in the computation of the Euler system. In [20], local truncation analysis is performed to prove that MPP flux limiters can maintain up to third order accuracy of the original scheme with no additional CFL constraint (i.e. $\text{CFL} \leq 1$) when the upwind flux is used. However, when the global LxF flux with extra large α in equation (4.1) is used, there is a mild time step restriction with $\text{CFL} \leq 0.886$. It is technically challenging to theoretically estimate such time step restriction for maintaining high order accuracy (e.g. fifth order) of the MPP flux limiters even for scalar equations, therefore we rely on extensive numerical tests.

We consider the scheme with the global Lax-Friedrichs flux with extra large $\alpha = 1.3$ (greater than $\max_u |f'(u)| = 1$). The time step is chosen to be $\Delta t = \text{CFL} \Delta x / \alpha$. In Table 4.2, we show that for the 5th order linear scheme (linear weights instead of nonlinear weights

Table 4.2: Example 4.2. Burgers' equation at $T = 0.2$. $\alpha = 1.3$ for the global LxF flux (4.1). $\Delta t = 0.886\Delta x/\alpha$. $u_{max} - v_{max}$ is the difference of the maximum values between the numerical solution and the exact solution.

	N	L^1 error	order	L^∞ error	order	$u_{max} - v_{max}$
without limiters	40	2.05E-04	—	2.76E-03	—	-5.47E-06
	80	1.20E-05	4.09	2.33E-04	3.57	-2.24E-07
	160	4.32E-07	4.79	9.68E-06	4.59	-1.98E-08
	320	1.38E-08	4.97	3.15E-07	4.94	-1.37E-09
	640	4.39E-10	4.98	1.01E-08	4.97	-8.92E-11
with limiters	40	2.06E-04	—	2.76E-03	—	7.33E-06
	80	1.20E-05	4.10	2.33E-04	3.57	9.99E-14
	160	4.32E-07	4.79	9.68E-06	4.59	1.00E-13
	320	1.38E-08	4.97	3.15E-07	4.94	1.00E-13
	640	4.39E-10	4.98	1.01E-08	4.97	9.99E-14

in WENO5) with the 4th order Runge-Kutta time discretization, when $CFL = 0.886$, the 5th order accuracy is maintained with the MPP flux limiters. In fact, $CFL = 0.886$ works for all other α 's we tested, the results are not listed here to save space.

Example 4.3. (Accuracy test for 2D vortex evolution problem.) We consider the vortex evolution problem [7] to test the accuracy. For this problem, the mean flow is $\rho = p = u = v = 1$ and is added by an isentropic vortex perturbation centered at (x_0, y_0) in (u, v) with $T = p/\rho$, no perturbation in entropy $S = p/\rho^\gamma$,

$$(\delta u, \delta v) = \frac{\varepsilon_{vortex}}{2\pi} e^{0.5(1-r^2)} (-\bar{y}, \bar{x}), \quad \delta T = -\frac{(\gamma-1)\epsilon^2}{8\gamma\pi^2} e^{(1-r^2)}, \quad \delta S = 0, \quad (4.2)$$

where $(\bar{x}, \bar{y}) = (x - x_0, y - y_0)$, $r^2 = \bar{x}^2 + \bar{y}^2$.

The computational domain is taken to be $[-5, 15] \times [-5, 15]$ and $(x_0, y_0) = (5, 5)$. The boundary condition is periodic. $\gamma = 1.4$ and the vortex strength is $\varepsilon_{vortex} = 10.0828$ as in [26]. The exact solution is the passive convection of the vortex with the mean flow. The lowest density and pressure of the exact solution are 7.8×10^{-15} and 1.7×10^{-20} .

ϵ_{WENO} in the nonlinear WENO weights is chosen to be 10^{-5} , which is between 10^{-2} and 10^{-6} [7]. In Table 4.3, we can clearly observe the 5th order accuracy with the PP flux limiters.

Example 4.4. 1D low density and low pressure problems. We consider two 1D low density and low pressure problems for the ideal gas. The first one is a 1D Riemann problem, the initial condition is $\rho_L = \rho_R = 7$, $u_L = -1$, $u_R = 1$, $p_L = p_R = 0.2$ and $\gamma = 1.4$, which is a

Table 4.3: Example 4.3. Vortex evolution problem at $T = 0.01$. $\epsilon_{WENO} = 10^{-5}$. ρ_{min} and p_{min} are the minimum density and pressure of the numerical solution respectively.

	N	L^1 error	order	L^∞ error	order	ρ_{min}	p_{min}
without limiters	64	1.49E-04	–	5.25E-02	–	-9.10E-05	2.79E-04
	128	1.57E-06	6.57	5.39E-04	6.61	3.04E-06	-2.16E-06
	256	1.29E-07	3.60	1.37E-04	1.97	-2.83E-06	-4.68E-07
	512	4.69E-09	4.79	3.37E-06	5.35	2.42E-07	1.27E-08
	1024	1.15E-10	5.35	7.92E-08	5.41	1.31E-08	1.87E-10
with limiters	64	1.49E-04	–	5.25E-02	–	6.30E-04	1.91E-04
	128	1.57E-06	6.57	5.39E-04	6.61	3.72E-06	1.00E-13
	256	1.32E-07	3.57	1.30E-04	2.05	2.42E-07	5.36E-10
	512	4.69E-09	4.81	3.37E-06	5.27	2.42E-07	1.27E-08
	1024	1.15E-10	5.35	7.92E-08	5.41	1.31E-08	1.87E-10

double rarefaction problem. The exact solution contains vacuum. In Fig. 4.1 (left), we show the results with the PP flux limiters at $T = 0.6$ on a mesh size of $\Delta x = 1/200$.

The second one is the 1D Sedov blast wave. For the initial condition, the density is 1, the velocity is 0, the total energy is 10^{-12} everywhere except in the center cell, which is a constant $E_0/\Delta x$ with $E_0 = 3200000$. $\gamma = 1.4$. In Fig. 4.1 (right), we show the results with the PP flux limiters at $T = 0.001$ on a mesh size of $\Delta x = 1/200$.

Example 4.5. 2D low density and low pressure problems. Now we consider two 2D low density and low pressure problems for the ideal gas. The first one is the 2D Sedov blast wave. The computational domain is a square of $[0, 1.1] \times [0, 1.1]$. For the initial condition, similar to the 1D case, the density is 1, the velocity is 0, the total energy is 10^{-12} everywhere except in the lower left corner is the constant $\frac{0.244816}{\Delta x \Delta y}$. $\gamma = 1.4$. The numerical boundary on the left and bottom edges is reflective. In Fig. 4.2 (left), we show the numerical density at the mesh sizes $\Delta x = \Delta y = \frac{1.1}{160}$ with the PP flux limiters at $T = 1$. The numerical solution with cutting along the diagonal matches the exact solution very well in Fig. 4.2 (right).

The second one is the shock diffraction problem. The computational domain is the union of $[0, 1] \times [6, 11]$ and $[1, 13] \times [0, 11]$. The initial condition is a pure right-moving shock of $Mach = 5.09$, initially located at $x = 0.5$ and $6 \leq y \leq 11$, moving into undisturbed air ahead of the shock. The undisturbed air has a density of 1.4 and a pressure of 1. The boundary conditions are inflow at $x = 0$, $6 \leq y \leq 11$, outflow at $x = 13$, $0 \leq y \leq 11$, $1 \leq x \leq 13$, $y = 0$ and $0 \leq x \leq 13$, $y = 11$, and reflective at the walls $0 \leq x \leq 1$, $y = 6$ and $x = 1$, $0 \leq y \leq 6$. $\gamma = 1.4$. The density and pressure at the mesh sizes $\Delta x = \Delta y = \frac{1}{32}$ with the PP flux limiters at $T = 2.3$ are presented in Fig. 4.3.

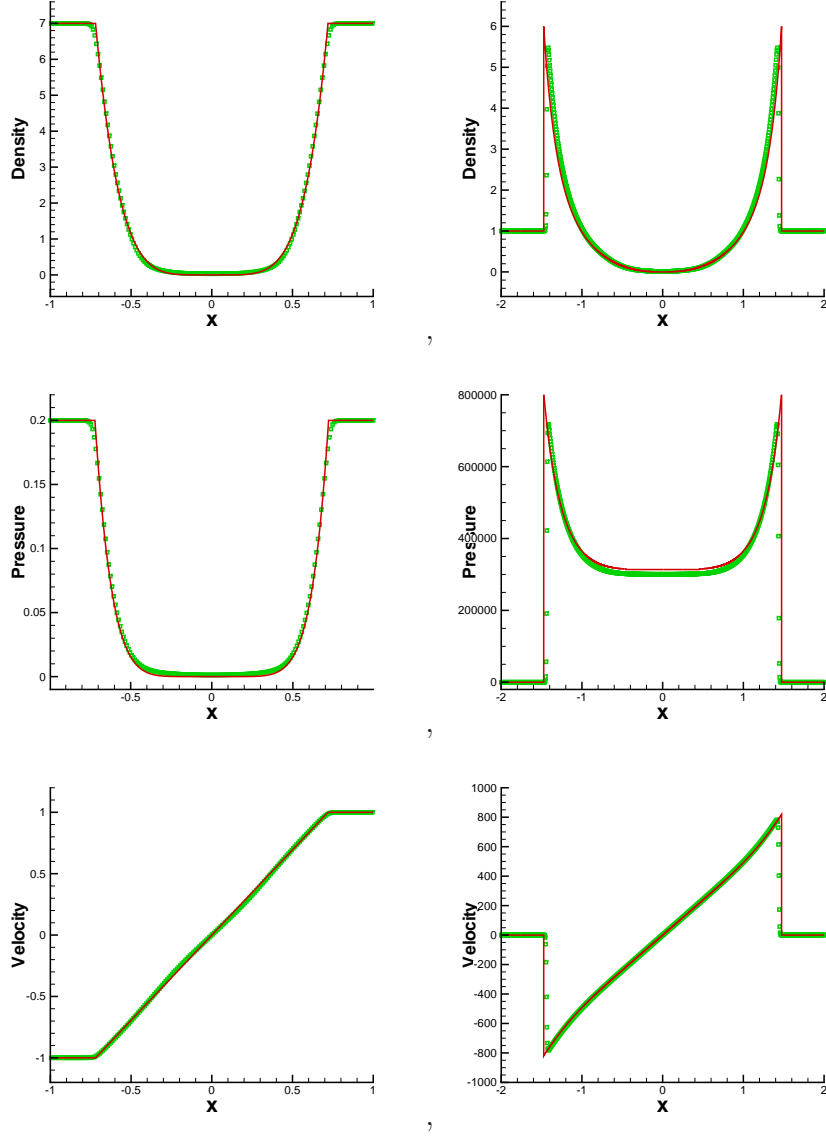


Figure 4.1: Example 4.4. Left: double rarefaction problem at $T = 0.6$. Right: 1D Sedov blast wave at $T = 0.001$. $\Delta x = \frac{1}{200}$. The solid lines are the exact solutions. Symbols are the numerical solutions.

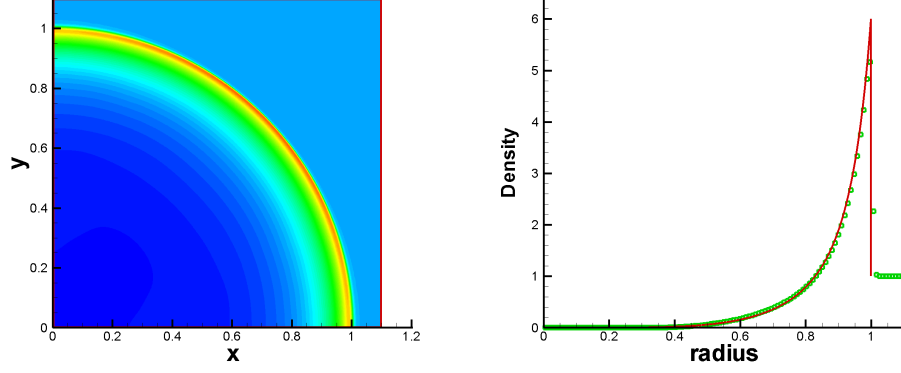


Figure 4.2: Example 4.5. 2D Sedov blast wave. $T = 1$. $\Delta x = \Delta y = \frac{1.1}{160}$. Left: contour of density. Right: cut along diagonal, the solid line is the exact solution, symbols are the numerical solution.

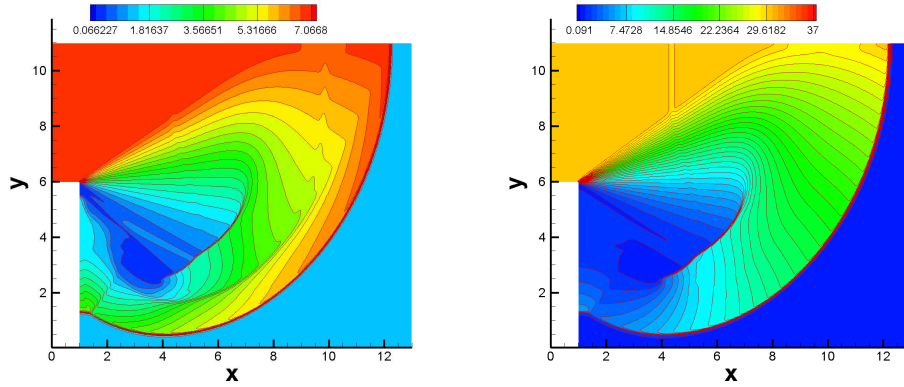


Figure 4.3: Example 4.5. 2D shock diffraction problem. $T = 2.3$. $\Delta x = \Delta y = \frac{1}{32}$. Left: density, 20 equally spaced contour lines from $\rho = 0.066227$ to $\rho = 7.0668$. Right: pressure, 40 equally spaced contour lines from $p = 0.091$ to $p = 37$.

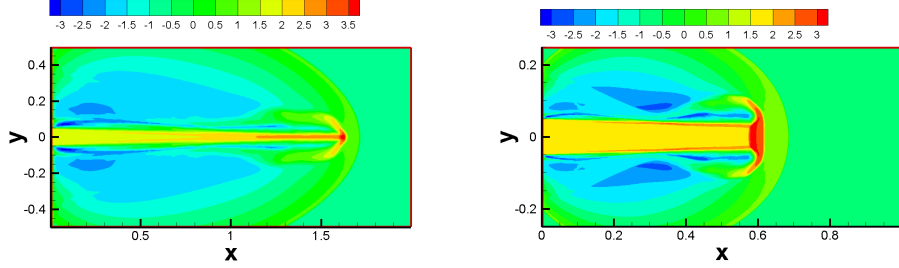


Figure 4.4: Example 4.6. High Mach number astrophysical jet. $T = 2.3$. Left: density of Mach 80 at $T = 0.07$ with mesh 448×224 . Right: density of Mach 2000 at $T = 0.001$ with mesh 800×400 .

Example 4.6. High Mach number astrophysical jets. We consider two high Mach number astrophysical jets without the radiative cooling [5, 26]. The first one is a Mach 80 problem. $\gamma = 5/3$. The computational domain is $[0, 2] \times [-0.5, 0.5]$, which is full of the ambient gas with $(\rho, u, v, p) = (0.5, 0, 0, 0.4127)$ initially. The boundary conditions for the right, top and bottom are outflows. For the left boundary, $(\rho, u, v, p) = (5, 30, 0, 0.4127)$ if $y \in [-0.05, 0.05]$ and $(\rho, u, v, p) = (5, 0, 0, 0.4127)$ otherwise. The numerical density on a mesh of 448×224 grid points with the PP flux limiters at $T = 0.07$ is shown in Fig. 4.4 (left). Then a Mach 2000 problem is considered to show the robustness of the scheme with the PP flux limiters. The computational domain is taken as $[0, 1] \times [-0.25, 0.25]$, initially full of the ambient gas with $(\rho, u, v, p) = (0.5, 0, 0, 0.4127)$. Similarly, the right, top and bottom boundary are outflows. For the left boundary, $(\rho, u, v, p) = (5, 800, 0, 0.4127)$ if $y \in [-0.05, 0.05]$ and $(\rho, u, v, p) = (5, 0, 0, 0.4127)$ otherwise. The numerical density at a mesh of 800×400 grid points with the PP flux limiters at $T = 0.001$ is shown in Fig. 4.4 (right).

Example 4.7. The reactive Euler equations. We consider the following two-dimensional Euler equations with a source term, which are often used to model the detonation waves [18, 26]:

$$\mathbf{u}_t + \mathbf{f}(\mathbf{u})_x + \mathbf{g}(\mathbf{u}) = \mathbf{s}(\mathbf{u}), \quad t \geq 0, \quad (x, y) \in \mathbb{R}^2, \quad (4.3)$$

$$\mathbf{u} = \begin{pmatrix} \rho \\ m_u \\ m_v \\ E \\ \rho Y \end{pmatrix}, \quad \mathbf{f}(\mathbf{u}) = \begin{pmatrix} m_u \\ \rho u^2 + p \\ \rho uv \\ (E + p)u \\ \rho u Y \end{pmatrix}, \quad \mathbf{g}(\mathbf{u}) = \begin{pmatrix} m_v \\ \rho uv \\ \rho v^2 + p \\ (E + p)v \\ \rho v Y \end{pmatrix}, \quad \mathbf{s}(\mathbf{u}) = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ \omega \end{pmatrix}, \quad (4.4)$$

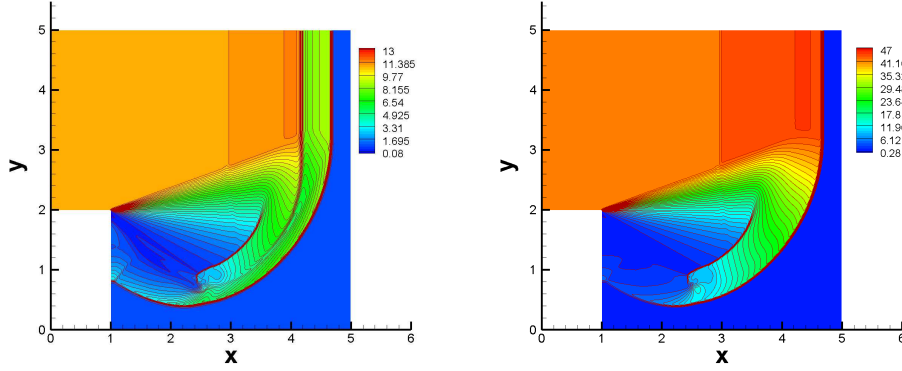


Figure 4.5: Example 4.7. Detonation diffraction at a 90° corner. $T = 0.6$. Mesh 400×400 . Left: density; Right: pressure.

with

$$m_u = \rho u, \quad m_v = \rho v, \quad E = \frac{1}{2}\rho u^2 + \frac{1}{2}\rho v^2 + \frac{p}{\gamma - 1} + \rho q Y,$$

where q is the heat release rate of reaction, γ is the specific heat ratio and Y is the reactant mass fraction. The source term is assumed to be in an Arrhenius form

$$\omega = -\tilde{K}\rho Y \exp(-\tilde{T}/T), \quad (4.5)$$

where $T = \frac{p}{\rho}$ is the temperature, \tilde{T} is the activation temperature and \tilde{K} is a constant. The eigenvalues of the Jacobian $\mathbf{f}'(\mathbf{u})$ are $u - c, u, u, u, u + c$ and the eigenvalues of the Jacobian $\mathbf{g}'(\mathbf{u})$ are $v - c, v, v, v, v + c$, where $c = \sqrt{\gamma \frac{p}{\rho}}$. The computation domain for this problem is the union of $[0, 1] \times [2, 5]$ and $[1, 5] \times [0, 5]$. The initial conditions are, if $x < 0.5$, $(\rho, u, v, E, Y) = (11, 6.18, 0, 970, 1)$; otherwise, $(\rho, u, v, E, Y) = (1, 0, 0, 55, 1)$. The boundary conditions are reflective except at $x = 0$, $(\rho, u, v, E, Y) = (11, 6.18, 0, 970, 1)$. Here the parameters are chosen to be $\gamma = 1.2$, $q = 50$, $\tilde{T} = 50$ and $\tilde{K} = 2566.4$.

This problem is similar to the shock diffraction problem in Example 4.5, but this one has a source term. The time step is taken to be

$$\Delta t = \frac{\text{CFL}}{\lambda_{\max}(\frac{1}{\Delta x} + \frac{1}{\Delta y}) + \tilde{K}}, \quad (4.6)$$

where $\lambda_{\max} = \max\{\|u\| + c, \|v\| + c\}$ on all grids, and \tilde{K} comes from the source term (4.5), such that the first order monotone scheme is PP. The numerical density and pressure at a mesh of 400×400 grid points with the PP flux limiters at $T = 0.6$ are shown in Fig. 4.5, which are comparable to the results in [18, 26].

Example 4.8. General equation of state. We consider the three species model of the one-dimensional Euler system with a more general equation of state in [19, 26]. The model involves three species, O_2 , O and N_2 ($\rho_1 = \rho_O$, $\rho_2 = \rho_{O_2}$ and $\rho_3 = \rho_{N_2}$) with the reaction



The governing equations are

$$\begin{pmatrix} \rho_1 \\ \rho_2 \\ \rho_3 \\ \rho u \\ E \end{pmatrix}_t + \begin{pmatrix} \rho_1 u \\ \rho_2 u \\ \rho_3 u \\ \rho u^2 + p \\ (E + p)u \end{pmatrix}_x = \begin{pmatrix} 2M_1\omega \\ -M_2\omega \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad (4.8)$$

and

$$\rho = \sum_{s=1}^3 \rho_s, \quad p = RT \sum_{s=1}^3 \frac{\rho_s}{M_s}, \quad E = \sum_{s=1}^3 \rho_s e_s(T) + \rho_1 h_1^0 + \frac{1}{2} \rho u^2, \quad (4.9)$$

where the enthalpy h_1^0 is a constant, R is the universal gas constant, M_s is the molar mass of species s , and the internal energy $e_s(T) = \frac{3RT}{2M_s}$ and $\frac{5RT}{2M_s}$ for monoatomic and diatomic species respectively. The rate of the chemical reaction is given by

$$\omega = \left(k_f(T) \frac{\rho_2}{M_2} - k_b(T) \left(\frac{\rho_1}{M_1} \right)^2 \right) \sum_{s=1}^3 \frac{\rho_s}{M_s}, \quad k_f = C_0 T^{-2} \exp(-E_0/T), \quad (4.10)$$

$$k_b = k_f / \exp(b_1 + b_2 \log z + b_3 z + b_4 z^2 + b_5 z^3), \quad z = 10000/T. \quad (4.11)$$

The parameters and constants are $h_1^0 = 1.558 \times 10^7$, $R = 8.31447215$, $C_0 = 2.9 \times 10^{17} m^3$, $E_0 = 59750K$, and $b_1 = 2.855$, $b_2 = 0.988$, $b_3 = -6.181$, $b_4 = -0.023$, $b_5 = -0.001$.

The eigenvalues of the Jacobian are $(u, u, u, u + c, u - c)$ where $c = \sqrt{\gamma \frac{p}{\rho}}$ with $\gamma = 1 + \frac{p}{T \sum_{s=1}^3 \rho_s e'_s(T)}$. Similar to Example 4.7, the time step is chosen to be

$$\Delta t = \frac{\text{CFL } \Delta x}{\lambda_{max} + s_{max} \Delta x}, \quad (4.12)$$

here $\lambda_{max} = \max\{|||u| + c||_\infty\}$ on all grids and s_{max} is

$$s_{max} = \max \left\{ \left| \frac{M_2 \omega}{\rho_2} \right|, \left| \frac{2M_1 \omega}{\rho_1} \right| \right\}. \quad (4.13)$$

A shock tube problem is considered for the reactive flows with high pressure on the left and low pressure on the right initially in the chemical equilibrium ($\omega = 0$). The initial conditions are:

$$(p_L, T_L) = (1000N/m^2, 8000K), \quad (p_R, T_R) = (1N/m^2, 8000K), \quad (4.14)$$

with zero velocity everywhere and the densities satisfying

$$\frac{\rho_1}{2M_1} + \frac{\rho_2}{M_2} = \frac{21}{79} \frac{\rho_3}{M_3}, \quad (4.15)$$

where $M_1 = 0.016$, $M_2 = 0.032$ and $M_3 = 0.028$. The initial densities of O , O_2 and N_2 are $5.251896311257204 \times 10^{-5}$, $3.748071704863518 \times 10^{-5}$ and $2.962489471973072 \times 10^{-4}$ on the left respectively, and $8.341661837019181 \times 10^{-8}$, $9.455418692098664 \times 10^{-11}$ and $2.748909430004963 \times 10^{-7}$ on the right respectively.

The numerical solution with the PP flux limiter is computed on a mesh size of $\Delta x = \frac{2}{4000}$ up to $T = 0.0001$. $\epsilon_{WENO} = 10^{-20}$ is taken as in [26]. In Fig. 4.6, the positivity of ρ_1 , ρ_2 , ρ_3 and p is preserved and converged solutions are observed.

5 Conclusion

We addressed the potential negative density and pressure problem that emerges when the high order WENO schemes are applied to solve compressible Euler equations in some extreme situations. The approach that we propose is in the conservative high order finite difference WENO approximation framework. We generalized the MPP flux limiting technique for the high order finite difference WENO methods solving scalar conservation law to a class of PP flux limiters for compressible Euler equations. We also developed the parametrized flux limiters for equations with source terms. Extensive numerical tests show the capability of the proposed approach: without sacrificing accuracy and much of the efficiency, the new schemes produce solutions satisfying the PP property for scalar problems with a source term, and solutions with positive density and pressure for compressible Euler equations with or without source terms.

Acknowledgement. We would like to thank Xiangxiong Zhang from MIT for helpful discussions.

References

- [1] J. P. BORIS AND D. L. BOOK, *Flux-corrected transport. I. SHASTA, A fluid transport algorithm that works*, Journal of computational physics, 11 (1973), pp. 38–69.
- [2] S. R. CHAKRAVARTHY AND S. OSHER, *High resolution applications of the Osher upwind scheme for the Euler equations*, AIAA, (1983).

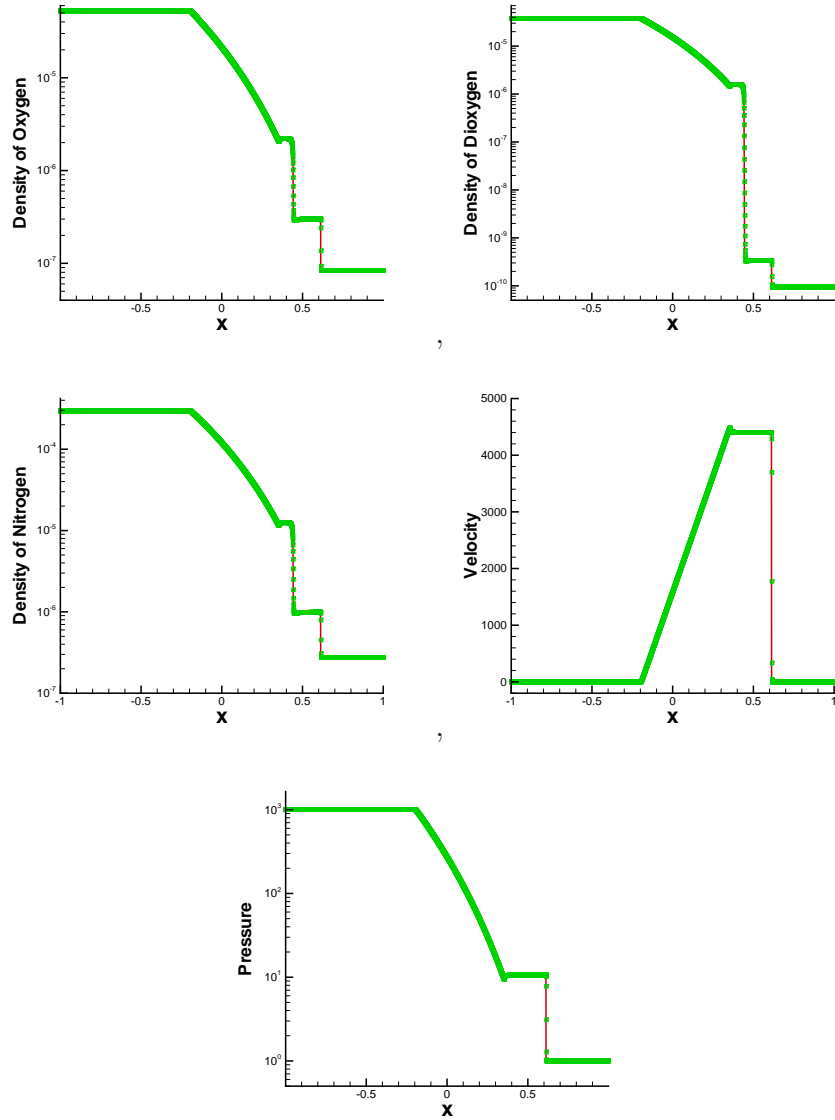


Figure 4.6: Example 4.8. Three species reaction problem at $T = 0.0001$. The solid lines are the reference solutions at $\Delta x = \frac{2}{8000}$. Symbols are the numerical solutions at $\Delta x = \frac{2}{4000}$.

- [3] B. EINFELDT, C.-D. MUNZ, P. L. ROE, AND B. SJÖGREEN, *On Godunov-type methods near low densities*, Journal of computational physics, 92 (1991), pp. 273–295.
- [4] B. ENGQUIST AND S. OSHER, *Stable and entropy satisfying approximations for transonic flow calculations*, Mathematics of Computation, 34 (1980), pp. 45–75.
- [5] Y. HA AND C. L. GARDNER, *Positive scheme numerical simulation of high mach number astrophysical jets*, Journal of Scientific Computing, 34 (2008), pp. 247–259.
- [6] A. HARTEN, B. ENGQUIST, S. OSHER, AND S. R. CHAKRAVARTHY, *Uniformly high order accurate essentially non-oscillatory schemes, III*, Journal of Computational Physics, 71 (1987), pp. 231–303.
- [7] C. HU AND C.-W. SHU, *Weighted essentially non-oscillatory schemes on triangular meshes*, Journal of Computational Physics, 150 (1999), pp. 97–127.
- [8] X. Y. HU, N. A. ADAMS, AND C.-W. SHU, *Positivity-preserving method for high-order conservative schemes solving compressible Euler equations*, Journal of Computational Physics, 242 (2013), pp. 169–180.
- [9] G.-S. JIANG AND C.-W. SHU, *Efficient implementation of weighted ENO schemes*, Journal of Computational Physics, 126 (1996), pp. 202–228.
- [10] C. LIANG AND Z. XU, *Parametrized maximum principle preserving flux limiters for high order schemes solving multi-dimensional scalar hyperbolic conservation laws*, Journal of Scientific Computing, (2013).
- [11] T. LINDE AND P. L. ROE, *Robust Euler codes*, in Thirteenth Computational Fluid Dynamics Conference, AIAA Paper-97-2098, 1997.
- [12] X.-D. LIU AND S. OSHER, *Nonoscillatory high order accurate self-similar maximum principle satisfying shock capturing schemes I*, SIAM Journal on Numerical Analysis, 33 (1996), pp. 760–779.
- [13] B. PERTHAME AND C.-W. SHU, *On positivity preserving finite volume schemes for Euler equations*, Numerische Mathematik, 73 (1996), pp. 119–130.
- [14] C.-W. SHU, *Essentially non-oscillatory and weighted essentially non-oscillatory schemes for hyperbolic conservation laws*, Advanced numerical approximation of non-linear hyperbolic equations, (1998), pp. 325–432.

- [15] C.-W. SHU AND S. OSHER, *Efficient implementation of essentially non-oscillatory shock-capturing schemes*, Journal of Computational Physics, 77 (1988), pp. 439–471.
- [16] P. K. SWEBY, *High resolution schemes using flux limiters for hyperbolic conservation laws*, SIAM journal on numerical analysis, 21 (1984), pp. 995–1011.
- [17] B. VAN LEER, *Towards the ultimate conservative difference scheme. II. monotonicity and conservation combined in a second-order scheme*, Journal of computational physics, 14 (1974), pp. 361–370.
- [18] C. WANG, X. ZHANG, C.-W. SHU, AND J. NING, *Robust high order discontinuous Galerkin schemes for two-dimensional gaseous detonations*, Journal of Computational Physics, 231 (2012), pp. 653–665.
- [19] W. WANG, C.-W. SHU, H. YEE, AND B. SJÖGREEN, *High-order well-balanced schemes and applications to non-equilibrium flow*, Journal of Computational Physics, 228 (2009), pp. 6682–6702.
- [20] T. XIONG, J.-M. QIU, AND Z. XU, *A parametrized maximum principle preserving flux limiter for finite difference RK-WENO schemes with applications in incompressible flows*, Journal of Computational Physics, 252 (2013), pp. 310–331.
- [21] T. XIONG, J.-M. QIU, Z. XU, AND A. CHRISTLIEB, *High order maximum principle preserving Semi-Lagrangian finite difference WENO schemes for the Vlasov equation*, arXiv:1401.0076, (submitted, 2013).
- [22] Z. XU, *Parametrized maximum principle preserving flux limiters for high order scheme solving hyperbolic conservation laws: one-dimensional scalar problem*, Mathematics of Computation, (in press.).
- [23] X. ZHANG AND C.-W. SHU, *On positivity-preserving high order discontinuous Galerkin schemes for compressible Euler equations on rectangular meshes*, Journal of Computational Physics, 229 (2010), pp. 8918–8934.
- [24] ———, *Maximum-principle-satisfying and positivity-preserving high-order schemes for conservation laws: survey and new developments*, Proceedings of the Royal Society A: Mathematical, Physical and Engineering Science, 467 (2011), pp. 2752–2776.

- [25] X. ZHANG AND C.-W. SHU, *Positivity-preserving high order discontinuous Galerkin schemes for compressible Euler equations with source terms*, Journal of Computational Physics, 230 (2011), pp. 1238–1248.
- [26] X. ZHANG AND C.-W. SHU, *Positivity-preserving high order finite difference WENO schemes for compressible Euler equations*, Journal of Computational Physics, 231 (2012), pp. 2245–2258.